

玉米全基因组谷胱苷肽 S 转移酶基因家族的分析

李晓玉, 江海波, 江海洋, 朱苏文

安徽农业大学生命科学学院, 安徽省作物生物学重点实验室, 合肥

摘要: 谷胱苷肽转移酶, 是植物生长发育过程中一类多功能酶, 在植物的初级代谢和次级代谢以及调控植物的逆境胁迫和细胞信号转导等过程中具有重要作用。根据 蛋白序列构建的隐马尔可夫模型, 鉴定玉米全基因组 GST 基因, 并对 GST 基因数量、类型、染色体定位和进化关系进行分析。结果表明, 玉米全基因组中共含有 个 GST 基因, 在 条染色体上呈非随机分布, 具有 型、型和 型 个亚族。与水稻比较 玉米未发现 亚族。该研究结果对深入了解 GST 基因家族的分子进化以及 GST 基因的克隆和功能验证提供了重要参考依据。

关键词: 玉米; 谷胱苷肽转移酶; 染色体定位; 进化

中图分类号:

文献标识码:

文章编号: - - -

Genome-wide analysis of glutathione-S-transferase gene family in maize

Abstract:

GST

GST

GST

GST

GST

Key words:

谷胱苷肽转移酶, 又称为谷胱苷肽 S 转移酶, 是催化谷胱苷肽与疏水的和亲电的化合物发生亲代取电反应的一种多基因家族编码的多功能蛋白酶, 广泛存在于动植物和微生物中。植物的 根据其基因结构和蛋白同源性的不同, 可以分为典型的 类, 包括 、 、 和脱氢抗坏血酸还原酶, 其中 和 是植物所特有的, 种类最多, 含量也最为丰富。

可溶性的 主要定位于细胞质中, 少数分布在叶绿体、微体, 也有少量存在于细胞核和质外体中。植物 最早在玉米中发现, 随后许多 GSTs 基因及类似序列相继从拟南芥、大豆、烟草及水稻等多个物种中分别被鉴定和克隆。

根据对目前已克隆的植物 GST 基因编码的蛋白质结构的分析, 植物 GSTs 基因存在一些共同的保守结构。 是一种球状二聚体蛋白一般以 ~ 的 条亚基以同源或异源的方式聚合而成,

收稿日期:

基金项目: 国家自然科学基金

和安徽省教育厅基金重点项目

共同资助。

作者简介: 李晓玉, 女, 博士, 讲师。

通信作者: 朱苏文, 女, 教授。

每个亚基都含有一个空间结构不同的基本结构域：一个是位于N端结构域，由一个氨基酸排列形成 β 折叠和一个 α 螺旋，是谷胱苷肽作用残基结合的活性位点。位点，保守性较高；二是由 \sim 个 α 螺旋构成的C端结构域，是亲电子物质的结合位点。而不同的在结构和功能上的差异主要体现在结构域上，不同种类的位点功能是不同的，位点结构的不同主要影响底物的特性。可溶性植物谷胱苷肽转移酶家族在进化树上呈显著分枝，每个亚族以希腊字母表示，如A类植物特有谷胱苷肽转移酶和B在进化树上构成大的分枝，而C和D转移酶构成很小的分枝。

在植物植物的生长和发育过程中具有重要的作用。目前的研究认为的主要生物学功能主要为以下几方面：去除外界毒素以及内源有毒代谢物的侵害。在植物体内能够催化某些内源或外来有害物质的亲电子基团与还原型谷胱苷肽的巯基结合，从而降低有害物质的毒性，例如：G类和H类的主要功能是与诸如除草剂、杀虫剂等在内的异生素结合，从而解除这些异生素毒性。作为结合蛋白起作用。早在1980年，等利用光亲和标记的方法，证明了是细胞分裂素和生长素的结合蛋白。提高植物的抗胁迫能力。胁迫因子可以使植物体内活性氧积聚，研究发现活性在胁迫条件下显著增加，加速清楚活性氧对植物的危害，说明在植物的抗氧化过程中具有十分重要的作用。与生长发育有密切关系。类型的GST基因的表达会受到细胞分裂素、生长素等激素的调控，在调控植物的生长和发育过程中具有重要作用。参与信号传导。研究认为通过蛋白-蛋白的相互作用参与信号传导，在信号传导中发挥作用，也是细胞程序性死亡的调节者。

玉米是我国重要的谷类作物，已成为研究进化生物学、遗传学和分子生物的模式植物。玉米基因组测序的完成使研究者能够在全基因组水平上对基因进行系统鉴定和比较基因组学的研究。尽管在植物的初级代谢、次级代谢、逆境胁迫和细胞信号转导等过程中具有重要作用，在玉米中仍没有系统的报道。因此，本研究通过生物信息学的方法对玉米全基因组GSTs基因进行鉴定和分析，研究结果将为玉米GSTs基因的克隆和功能验证奠定基础。

1 材料与方法

1.1 玉米基因组数据库序列来源

玉米基因组数据库来源于NCBI，使用的是最新版精细基因组数据。

1.2 玉米GST基因的确定

为了鉴定玉米全基因组中的GST基因，在一些可应用的公共数据库中进行了一系列的搜索与查询。首先，在NCBI网站、EMBL网站和DDBJ网站上，以“*GST*”，“*GST1*”或者“*GST2*”为关键词进行搜索来查询可能的玉米GST基因；第二步，将在网站上查询到的蛋白序列递交到

蛋白家族数据库中进行一系列分析，获得了N端结构域和C端结构域的氨基酸序列，由于N端结构域存在于迄今为止研究过的所有GST基因中，并且它是高度保守的，所以以N端结构域的氨基酸序列为标准序列，与玉米GST1进行同源搜索；第三步，使用ClustalW软件上来进行多序列比对程序进行多序列比对，去除冗余序列；第四步，将得到的非重复候选基因序列分别通过NCBI数据库服务系统进行分析，去除其中不含有N端结构域的序列。

1.3 玉米GST基因的命名和分类

根据Suzuki等提出的方法和原理对玉米GST基因进行命名和分类。对已命名的玉米GST基因，使用已命名的名字。对于没有研究过的玉米GST基因的命名和分类，采用与其它已研究的植物家族构建进化树，并根据进化关系进行确定。第一部分表示植物名，用相应植物拉丁名的属名和种加词的前几位字母缩写来表示；第二部分表示GST家族、亚族及成员编号。例如：水稻GST家族A亚族第1个基因表示为*OsGSTF1*。按照这种命名方法，玉米GST家族A亚族第1组的第1个基因表示为*ZmGSTT1*，其他GST家族的命名依次进行类推。

1.4 玉米GST基因的序列特征分析

为分析玉米GST基因的氨基酸序列属性，使用在线工具对获得的玉米GST基因编码的氨基酸序列进行等电点和蛋白分子量属性的计算。

1.5 玉米 GST 基因进化树的构建

使用 软件提供的 工具对玉米所有的 GST 基因的蛋白序列进行多序列比对, 然后进行邻接法

进行系统进化树的构建, 从而分析玉米 GSTs 基因家族的进化关系。同时使用玉米和水稻蛋白序列构建了复合进化树, 根据进化树分枝远近程度的不同, 对玉米 GST 基因进行分类。

1.6 GST 基因的染色体物理定位

玉米 GST 基因通过玉米基因组数据库提供的染色体 工具进行玉米染色体物理定位。首先将每个 GST 基因与网站公布的玉米基因组进行 , 获得每个 GST 基因在染色体上的起始位置, 所有 GST 基因使用这种方法确认位置后, 使用染色体定位绘图软件

, 将玉米 GST 基因族在玉米染色体上的分布与位置以图形的形式直观地表示出来。

2 结果与分析

2.1 玉米 GST 基因的确定、分类及命名

通过查询, 将所得玉米蛋白序列提交到蛋白家族数据库进行数据分析, 获得 结构域序列, 并作为标准序列与玉米数据库进行同源比对, 初步筛选出 个玉米 GST 基因。运用 软件上进行多序列比对程序, 删除冗余序列以及 和 的验证, 最终确定了 个玉米谷胱苷肽转移酶基因。同时, 通过对玉米 GST 基因的核苷酸数目、分子量, 以及氨基酸等电点进行了分析, 发现玉米 家族蛋白质分子量介于 和 之间表 。通过对 个玉米和 个水稻 蛋白序列构建的复合进化树, 将玉米 GST 基因分成了 个亚族, 包括 个 型、 个 型和 个 型。

2.2 玉米 GST 基因的染色体定位

图 展示了 个 GST 基因在玉米 条染色体上的分布, 可以明显地看出这些基因在染色体上分布不均匀。在 号染色体上, GSTU 基因分布数量最多, 一共有 个 GSTU 基因, 个基因定位在 号染色体, 个基因定位在 号染色体, 个基因定位在 号染色体, 个基因定位在 号染色体, 个基因定位在 号染色体, 而 号和 号染色体上则没有 GST 基因。

在玉米中, 亚族的 GST 基因分布在一个染色体上, 数目都在一个以上。但是有趣的是, 所有亚族 GST 基因均定位在 号染色体上, 且它们中的一个序列高度重叠相似, 并且构成了一个基因簇。根据 的定义 基因簇是在 的核苷酸单位中含有 一个以上基因的基因群 可以得出在玉米 GST 基因中共 个基因簇不均衡的分布在玉米的 条染色体上, 除了 号染色体含有 个基因簇分别包括 个基因外, 其余 条染色体 , , 各含有 个基因簇, 分别包括 个、 个和 个基因。

2.3 玉米 GST 基因的进化分析

为了进一步研究玉米 GST 基因家族内部基因间的进化关系, 根据玉米 蛋白序列构建进化树。如图 所示, 个玉米 GST 基因被分为 个亚族, 进化关系零散, 亚族分布在 个分支中, 被 、 亚族隔断; 同时 亚族分布在 个分支中, 被 、 亚族所隔断; 亚族与 分布在 个分支。 、 亚族均与 亚族中的一部分基因排列紧凑, 进化关系亲近, 但 、 内部某些基因之间进化关系较远。所有的 和 亚族都含有 和 结构域, 的 亚族, 的 亚族和 的 亚族含有 和 结构域。

2.4 玉米和水稻 GST 基因的比较进化分析

为研究玉米和水稻 GST 基因的进化关系, 构建了 个包含 个 蛋白序列的进化树, 包括 个玉米 序列和 个水稻 蛋白序列。在进化树中, 玉米和水稻 GST 基因团簇成 组, 包括 、 、 和 亚族。在玉米和水稻 GST 家族中, 除了 亚族之外, 其他亚族为共同所有, 水稻中含有 亚族, 而玉米中没有发现这个亚族。在第 分支中, 玉米 GSTU9, GSTU10, GSTU11, GSTU12 和 GSTU13 与水稻中的一些 GST 基因紧密团簇一起; 在第 分支中, 水稻 亚族, 亚族, 与玉米 亚族, 亚族混合团簇在一起; 在第 、第 分支中, 水稻 亚族和玉米 亚族又紧密排列在一起; 在第 分支中水稻 亚族与玉米 亚族团簇在一起。有趣的是, 玉米水稻 亚族与 亚族在进化树中交替团簇排列, 表明水稻与玉米 亚族和 亚族在长期进化机制比较复杂图 。

从图 可以看出, 玉米和水稻中的进化关系比较零散, 分支较多, 出现了多级分枝现象, 并且同一亚族集中分布各级分枝中, 出现被其他亚族隔断的现象。例如, 亚族被 亚族隔断。值得注意的是, 在第 分支中, 玉米 亚族与 亚族分布在

个分支中,说明在这些基因中, 亚族与 亚族进化关系紧密。

通过比对发现,在一个物种的比较进化树中出现了 对水稻 水稻之间亲源关系相近的基因对,例如: *OsGST19* 和 *OsGST22*; 对玉米 玉米之间的

亲源关系相近对,例如: *ZmGST16* 和 *ZmOsGST17*; 而水稻 玉米配对的亲源关系基因对仅有 对,例如: *ZmGSTF7* 和 *OsGSTF3*,可见这一个物种的 *GST* 基因在结构上有较大的差异性,其进化亲缘关系也相差较大。

表 1 玉米 37 个谷胱苷肽转移酶基因的序列特征

GST						
基因命名	编号	染色体	染色体位置	氨基酸长度	分子量	等电点
<i>ZmGSTU1</i>						
<i>ZmGSTU2</i>						
<i>ZmGSTU3</i>						
<i>ZmGSTU4</i>						
<i>ZmGSTU5</i>						
<i>ZmGSTU6</i>						
<i>ZmGSTU7</i>						
<i>ZmGSTU8</i>						
<i>ZmGSTU9</i>						
<i>ZmGSTU10</i>						
<i>ZmGSTU11</i>						
<i>ZmGSTU12</i>						
<i>ZmGSTU13</i>						
<i>ZmGSTU14</i>						
<i>ZmGSTU15</i>						
<i>ZmGSTU16</i>						
<i>ZmGSTU17</i>						
<i>ZmGSTU18</i>						
<i>ZmGSTU19</i>						
<i>ZmGSTU20</i>						
<i>ZmGSTU21</i>						
<i>ZmGSTU22</i>						
<i>ZmGSTU23</i>						
<i>ZmGSTU24</i>						
<i>ZmGSTU25</i>						
<i>ZmGSTU26</i>						
<i>ZmGSTU27</i>						
<i>ZmGSTF1</i>						
<i>ZmGSTF2</i>						
<i>ZmGSTF3</i>						
<i>ZmGSTF4</i>						
<i>ZmGSTF5</i>						
<i>ZmGSTF6</i>						
<i>ZmGSTF7</i>						
<i>ZmGSTZ1</i>						
<i>ZmGSTZ2</i>						
<i>ZmGSTZ3</i>						

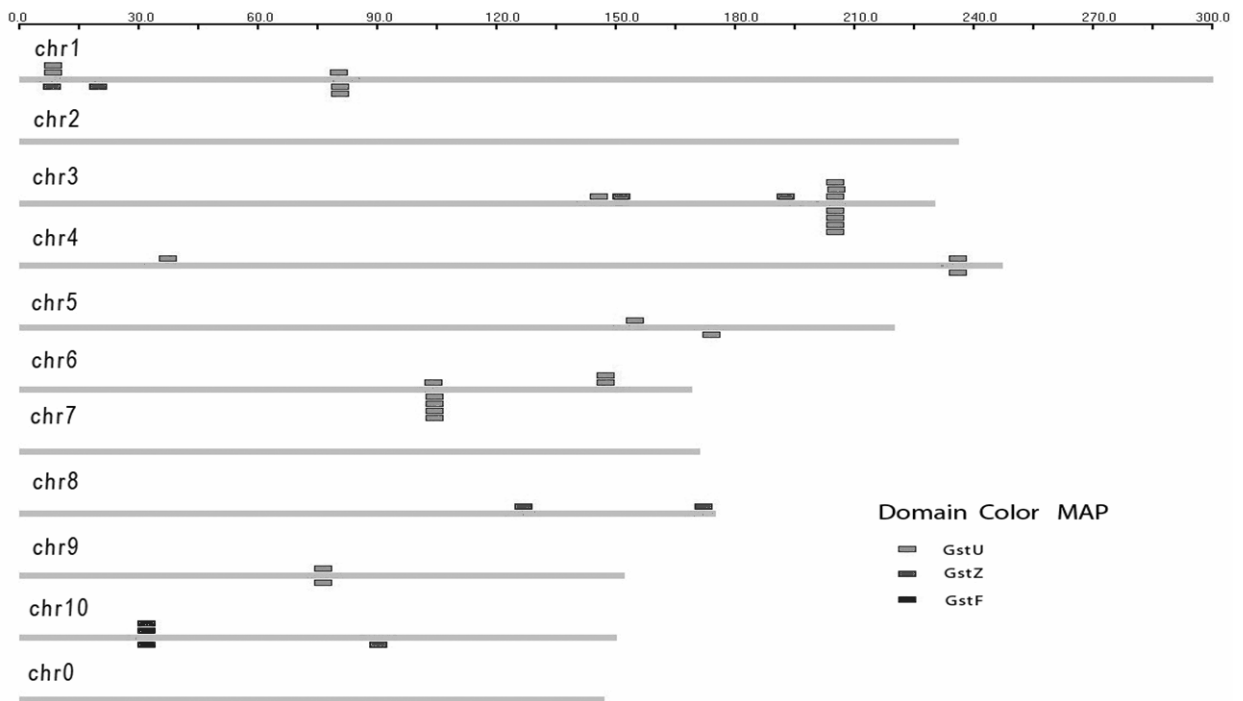


图 1 玉米谷胱苷肽转移酶的染色体定位
GST

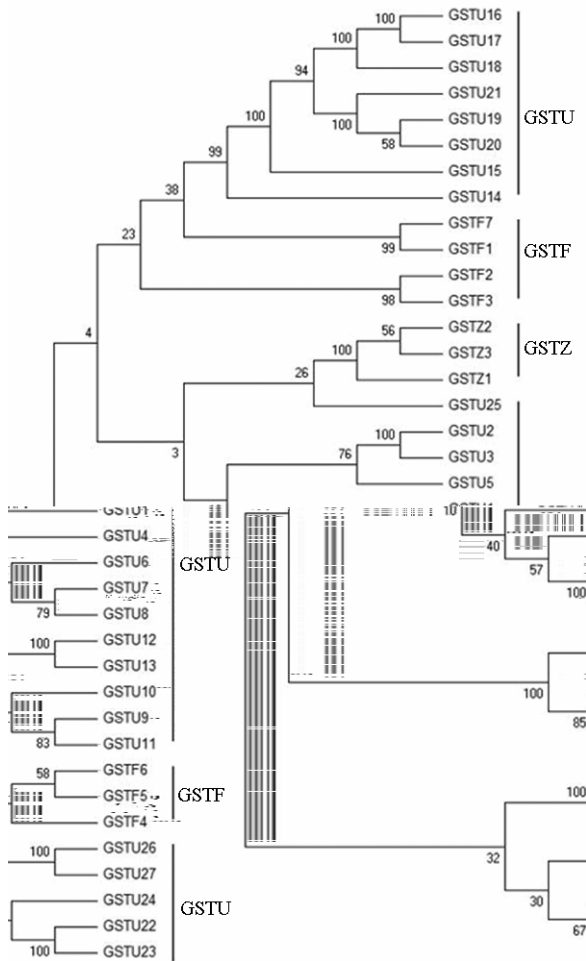


图 2 玉米 GST 进化关系分析

GST

3 讨论

等人在拟南芥基因组组中发现了一个谷胱苷肽转移酶，这些基因被分为、、、、、和亚族，等人在水稻中发现一个谷胱苷肽转移酶，这些基因被分为、、和亚族。本研究将结构域序列作为标准序列，与玉米基因组本地库进行本地比对，经过删除重复序列，最终在玉米中确定了一个谷胱苷肽转移酶基因。通过与水稻谷胱苷肽转移酶建立进化树进行进化关系比较的方法，玉米家族被分为个亚族，分别为、和亚族。与动物相比，植物谷胱苷肽转移酶中含有一个特有的亚族，和亚族，而这些基因在动物中是没有的。可能的解释为，高等植物细胞谷胱苷肽转移酶与动物相比具有独特的功能，除了在除草剂解毒作用之外，通过改变表达量，植物胱苷肽转移酶也对生物、非生物胁迫，激素和发育改变的反应有关。

对于植物特有的亚族，和亚族，经历了集中的复制在万年前。在这个亚族中发生大量复制的原因尚不可知，可能的解释是不同的谷胱苷肽转移酶亚族有不同的特定功能，在至今发现的人类和线虫谷胱苷肽转移酶中亚族蛋白同源性为，人类和康乃馨谷胱苷

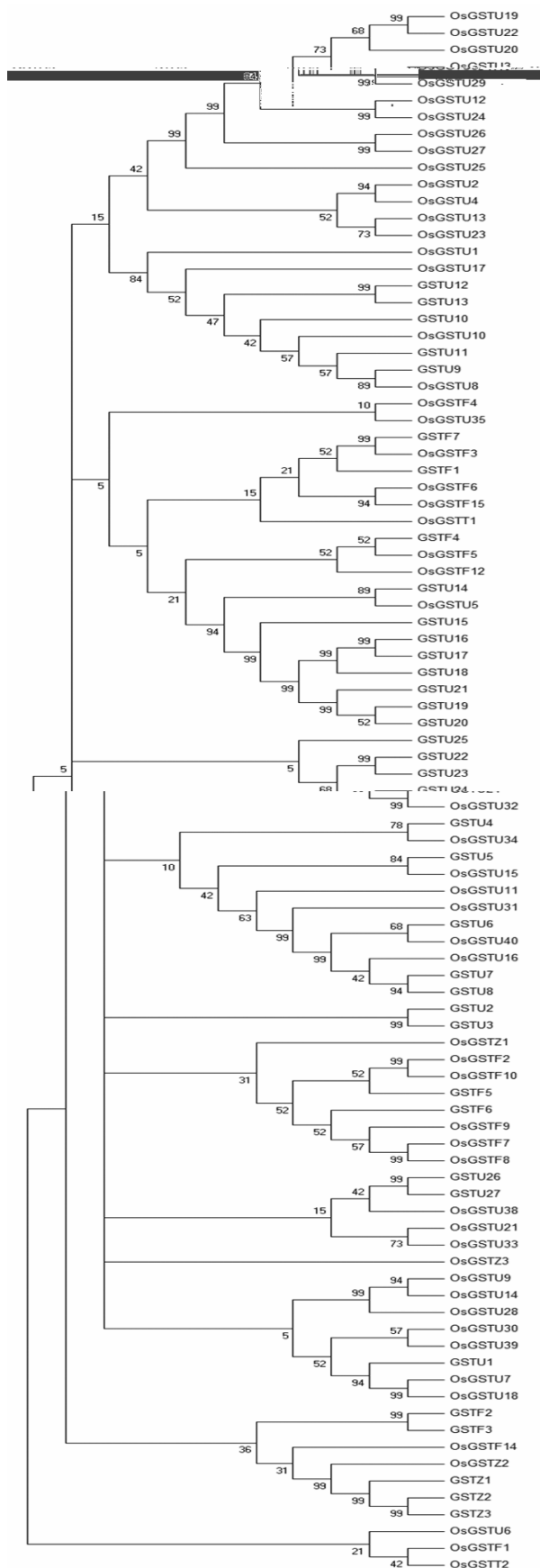


图 3 玉米和水稻谷胱苷肽转移酶基因进化关系分析

GST

肽转移酶中蛋白同源性为 $0.15 \sim 0.99$ ，表明一个共同的生物功能也许要经历长期的进化历程。但是在玉米谷胱苷肽转移酶中并没有发现 γ 亚族，可能的原因是在玉米 γ 和 δ 亚族长期复制过程中， γ 亚族渐渐缺失，并且玉米 δ 亚族除了与水稻中 δ 亚族中相同的功能之外，还具有水稻中 δ 亚族相对应的功能。与水稻 δ 进化树亚族独立成支相比，通过比邻法构建的玉米谷胱苷肽转移酶进化树中，形成 1 个比较独立的分支， δ 与 δ 、 δ 、 δ 、 δ 团簇成一个分支； δ 与 δ 、 δ 、 δ 组成一个较小的分支，剩下的 δ 亚族与 δ 亚族构成最大的一个分支。这些现象表明在玉米 δ 亚族中有些成员功能可能与 δ 亚族或者 δ 亚族成员类似。

对玉米谷胱苷肽转移酶的染色体定位研究显示 GST 基因在玉米染色体上的分布极不均匀，一共有 1 个 GST 基因定位在 1 号染色体上，为 GST 基因分布数量最多的染色体，但是在 2 号和 3 号染色体上没有发现任何 GST 基因。与玉米中其他 GST 亚族基因贯穿分布在许多条染色体上不同，所有的 δ 亚族成员集中分布在 1 号染色体上。因此，推测玉米中 δ 亚族比其他 δ 亚族的产生时期要晚。值得注意的是， δ 亚族作为玉米谷胱苷肽转移酶成员最多的一类，其在染色体上形成基因簇的现象也是最多的，共有 10 组基因簇，而 δ 亚族却没有发生基因簇现象，其原因和规律有待于进一步研究。

参考文献：

Arabidopsis thaliana

Arabidopsis thaliana

冯雪 王彬 孙艳香. 谷胱甘肽硫转移酶的研究进展
生物学教学

聂立红 王声湧 胡毅玲 谷胱甘肽硫 转移酶研究进
展 中国病理生理杂志

杨海灵 聂力嘉 朱圣庚 等. 谷胱甘肽硫转移酶结构
与功能研究进展 成都大学学报

Arabidopsis thaliana

刘新仿 李家洋 紫外线强烈诱导的谷胱甘肽转移酶
基因的功能鉴定 遗传学报

Arabidopsis thaliana